

PCAS 20 (2017/2018)

**Supervised Project Report
(ANTA604)**

***Tell Me About Antarctica:
Guidelines for In Situ Capture and Viewing of 360-
degree Video Highlighting Antarctic Science***

Rob Lindeman

Student ID: 89241433

Word count: 4,049 (excluding abstract and references)

Abstract/executive summary:

This report describes new methods for capturing 360-degree video content of Antarctic experts describing their work from Antarctica. In addition, the report describes the use of virtual reality (VR) techniques for creating engaging viewing experiences for consumers of the captured video. Several videos were captured in Christchurch of subject matter experts talking about their work, and lessons learned from these capture sessions were used in designing further capture sessions carried out in Antarctica.

The report includes links to viewable examples of the various concepts and techniques developed for this project, and ends by providing a set of new camera-shot types that uniquely apply to the new medium of 360-degree video.

Tell Me About Antarctica: Guidelines for In Situ Capture and Viewing of 360-degree Video Highlighting Antarctic Science

Robert W. Lindeman

Introduction

The goal of this project was to identify ways in which Antarctic scientists could effectively convey their research to the general public through personal digital media, focusing on creating first-person learning experiences with 360-degree videos and immersive VR viewers.

The motivation for this work came from my own experience during certain guest lectures we experienced as PCAS students. The bulk of the material presented during the PCAS Programme involved subject matter experts speaking to the class about their areas of research. Sometimes this involved classroom lectures, other times short field trips. On several of these occasions, I was struck by the passion exhibited by the speaker about their given topic, and how contagious this passion was. I started to wonder if there was a way to capture that passion and present it to those who could not take part in the PCAS Programme. Later, I pondered whether there might be some way to expand this beyond PCAS material to a more-general framework for the presentation of scientific wonder to the general public.

As a long-time virtual reality (VR) researcher, I understood the powerful effect that VR can have on people, and wondered if there was a way to marry these two topics. I had only viewed a small number of 360-degree videos prior to this course, and had always seen this medium as a poor stepchild of fully interactive VR. The “interaction” was limited to moving the head around, while “real” VR allowed the user to use the hands, feet, etc. to interact. The more I thought about it, however, the more I began to think that immersive 360-degree video might be able to bridge the gap between passive traditional video and fully interactive VR, and provide deeper engagement for learners. This led me to the idea of exploring how this technology might be used as an accessible way to convey scientific concepts in an engaging manner.

Method

I used 360-degree video and audio to individually record subject matter experts (SMEs) talking about their respective fields of expertise. Viewers will be able to view the content using various devices, such as a mobile phone, standard desktop monitor, and/or a VR headset.

After experimentation, it was found that the main challenges for the use of 360-degree video were selecting the correct point of view for capturing the content, and whether to capture using a hand-held (selfie-stick) or fixed (tripod) camera support. In addition, the question of whether and how to move the camera during capture was also found to be problematic. These three issues will be systematically addressed in this report, providing a set of guidelines for effective capture of 360 video for conveying information from Antarctic SMEs speaking in the field.

We first conducted several capture sessions of SMEs (Paul Broady and Colin Monteath) in Christchurch in order to explore some of the issues in a less-extreme environment, and to allow us to make several takes to vary some of the parameters. Following this, we then captured actual SMEs (Chris Long, Emma Beech, Jonathan Taylor) in Antarctica, discussing topics in the field.

Related Work

I conducted a literature search for the main ideas behind the proposed research, namely the idea that immersive technologies could improve engagement and learning, and that how the content was captured would make a difference in the effectiveness of presenters in their efforts.

360-Degree Video

The notion of 360-degree video (also called “immersive video” or “spherical video”) refers to the capture of video images from a set of cameras (or camera lenses) rigidly connected to a single point in space, and facing in enough directions to capture the entire scene simultaneously. The captured images are then computationally processed, taking into account the locations, orientations and properties of the cameras to create a spherical image. The resulting spherical image can then be viewed using either standard display hardware (e.g., a phone, tablet or desktop) or a special stereoscope-type device, such as a Google Cardboard¹ (Figure 1) or Samsung GearVR², so called “VR viewers”. Interacting (i.e., panning and tilting) with the spherical image on tradition displays using either a mouse or finger swipes allows the viewer to look in any direction from the captured location. However, interacting with VR viewers is accomplished by simply turning or tilting the head, providing a more-natural way of interacting.



Figure 1: User with Google Cardboard (Image: Creative Commons)

While both methods of viewing have their strengths and weaknesses, VR viewers are considered more immersive, in that the captured content becomes the primary focus of the viewer, and the head-movement interaction removes the need to handle the device through swiping or a mouse.

¹ <https://vr.google.com/cardboard/>

² <http://www.samsung.com/global/galaxy/gear-vr/>

VR and Immersion

There are several ways to categorise the technologies we will be using here. Slater, Usoh and Steed (1994) make a differentiation between *immersion* and *presence*. Immersion, they argue, is an *objective* description of the technology used to convey the content. By this definition, a standard desktop monitor would be less immersive than a surround projection system, since the former has a lower field of view. By contrast, *presence* is defined as “a psychological emergent property of an immersive system, and refers to the participant's sense of 'being there' in the world created by the [virtual reality] system.” They go on to claim that “immersion is a necessary rather than a sufficient condition for presence - immersion describes a kind of technology, and presence describes an associated state of consciousness.” Slater et al. also promote the use of “body-centred interaction,” which means mapping physical user movements to changes in the virtual world in a natural (expected) way. Most virtual reality researchers of immersion and presence since this seminal paper was published have adopted similar definitions of immersion and presence.

From the learning sciences, Chris Dede (2009) uses the term immersion in a slightly different way. He describes the notion of immersion as “the subjective impression that one is participating in a comprehensive, realistic experience”, which involves “the willing suspension of disbelief.” This definition seems to blur the line between immersion and presence, and underscores the difficulty in classifying immersive experiences.

Dede (2009) goes on to state that the “design of immersive learning experiences that induce this disbelief draws on sensory, actional, and symbolic factors.” One could map Dede's “sensory” immersion factors to Slater et al.'s (1994) overall definition of immersion. Dede's “actional” factors refer to the idea that a person feels more present in an experience when she can exhibit greater control over the experience, such as altering the point of view based on head rotation (looking left), which maps well to Slater et al.'s body-centred interaction idea. In terms of “symbolic” factors, Dede promotes the use of content that is grounded in the familiar world of the participant. He claims that “[i]nvoking digital versions of archetypical situations from one's culture deepens the immersive experience by drawing on the participant's beliefs, emotions, and values about the real world.”

Based on this previous work, we think that a virtual reality headset should be the preferred method for delivering the content of 360-degree educational videos, since it will be more immersive, and will support a natural way of controlling what the person sees.

Getting the Right Shot

In filmmaking for traditional cinema, there is a relatively established “lexicon” of camera shots or movements. These are used in various ways to achieve the desired effects. Some filmmakers are well known for their preference for specific types of shots, and some genre (e.g., Westerns) are also known to employ certain shot types more than others (e.g., “Cowboy” for Westerns). Table 1 lists 30 of the most popular shot types used in traditional cinema.

Table 1: Thirty Common Traditional Camera Shot Types³

#	Shot Name	Description
1	Aerial Shot	An exterior shot filmed from the air. Often used to establish a (usually exotic) location.
2	Arc Shot	A shot in which the subject is circled by the camera.
3	Bridging Shot	A shot that denotes a shift in time or place, like a line moving across an animated map.
4	Close Up	A shot that keeps only the face full in the frame.
5	Cowboy Shot	A shot framed from mid thigh up (including the holster), so called because of its recurrent use in Westerns.
6	Crane Shot	A shot where the camera is placed on a crane or jib and moved up or down.
7	Deep Focus	A shot that keeps the foreground, middleground and background ALL in sharp focus.
8	Dolly Zoom	A shot that sees the camera track forward toward a subject while simultaneously zooming out creating a woozy, vertigo effect.
9	Dutch Tilt	A shot where the camera is tilted slightly on its side to create a kooky angle.
10	Establishing Shot	A shot, at the head of the scene, that clearly shows the locale the action is set in.
11	Handheld Shot	A shot in which the camera operator holds the camera during motion to create a jerky, immediate feel.
12	High Angle Shot	A shot looking down on a character or subject, often isolating them in the frame. This can make the subject look meek.
13	Library Shot	A pre-existing shot of a location, typically a wild animal, that is pulled from a library.
14	Locked-Down Shot	A shot where the camera is fixed in one position while the action continues off-screen.
15	Long Shot	A shot that depicts an entire character or object from head to foot.
16	Low Angle Shot	A shot looking up at a character or subject often making them look bigger in the frame. It can make everyone look heroic and/or dominant.
17	Matte Shot	A shot that incorporates foreground action with a background, traditionally painted onto glass, now created in a computer.
18	Medium Shot	The shot that utilises the most common framing in movies, shows less than a long shot, more than a close-up.

³ <https://www.empireonline.com/movies/features/film-studies-101-camera-shots-styles/>

19	Money Shot	A shot that is expensive to shoot, but deemed worth it for its potential to wow, startle and generate interest.
20	Over-The-Shoulder Shot	A shot where the camera is positioned behind one subject's shoulder, usually during a conversation.
21	Pan	A shot where the camera moves continuously right to left or left to right.
22	POV shot	A shot that depicts the point of view of a character so that we see exactly what they see.
23	The Sequence Shot	A long shot that covers a scene in its entirety in one continuous sweep without editing.
24	Steadicam Shot	A shot from a hydraulically balanced camera that allows for a smooth, fluid movement.
25	Tilt	A shot where the camera moves continuously Up to Down or Down To Up.
26	Top Shot (Bird's-Eye Shot)	A shot looking directly down on a scene rather than at an angle.
27	Tracking Shot	A shot that follows a subject, be it from behind or alongside or in front of the subject.
28	Two-Shot	A medium shot that depicts two people in the frame.
29	Whip Pan	A shot that is the same as a pan, but is so fast that picture blurs beyond recognition.
30	Zoom	A shot deploying a lens with a variable focal length that allows the cinematographer to change the distance between camera and object without physically moving the camera.

This work provided me with a starting point from which to design my approach to capturing scientists in ways that might be effective.

Summary

From these sections, it was decided that the work should focus on viewing using an immersive VR viewer, and to use traditional camera shot descriptions as a starting point for capturing video content using 360-degree video. The next step was to prepare for capturing the first footage for the project.

Preparation for Shooting

A comparison was made between traditional video recording and 360-degree video recording, in order to tease out the essence of each. Then, decisions were made about which traditional shot types might be most appropriate, and which ones least.

Camera Hardware

The camera used in capturing all of the footage for this report was the Ricoh Theta V⁴. The camera has the advantage of ease of use, as it captures two fish-eye images and blends them into a single spherical image right on the camera. It also provides fairly advanced image capture options, such as settings for different lighting conditions and exposure times. In addition, it can capture both still and video footage, and be controlled remotely using a standard mobile phone. It also worked well in the cold. The negative side is that the memory is built in, with no option for expansion. This limited the video I could capture in the field, and forced me to trade off image resolution for the amount of footage.

360-degree Video Content vs. Traditional Video Content

There are several challenges in creating effective 360-degree videos compared to traditional videos. One is the lack of knowledge about what the position of the viewer will be when watching the content. For traditional video, the assumption is that the person is seated. But in the end, this may not matter, since the content does not change based on the posture of the viewer as it does with 360-degree video viewed with a VR viewer. For 360-degree video, the viewer may be seated in a stationary or swivel chair, standing, lying down, etc., and viewer posture matters, since head orientation is sensed to control the viewing direction on VR viewers.

A larger problem has to do with how to guide the viewer to see the action taking place in the video sequence. Because the viewer is free to look around, it could be that she misses important content, because she is looking in a different direction. This problem still has yet to be solved in the general case, and is beyond the scope of the current work. Since our target content has been selected to be educational video presented by a single presenter, we assume the viewer is looking at the presenter, and follows normal social conversational cues during the experience, such as looking where the presenter is indicating, focusing on things the presenter is talking about, and otherwise paying attention to the presenter. While these assumptions cannot be guaranteed, they are realistic for the target audience (i.e., people interested in the subject).

The list of shots for traditional film shown in Table 1, though not exhaustive, supports the artistic creativity necessary for film-based expression, across multiple genre and story elements. When deciding how to capture the SMEs for this new medium of 360-degree video, we first thought about how to use traditional shot techniques. The ability to look completely around in a scene, however, makes some of the shots (e.g., Pan, Two-Shot, Locked-Down Shot) unnecessary. Also, since the video will be viewed from a first-person viewpoint, techniques such as the Low Angle or High Angle shots might artificially influence the viewer into feeling subservient or superior to the presenter. Additionally, as mentioned earlier, large or sudden movements in VR have been shown to induce motion sickness in users, so many of the camera movement techniques that involve sudden involuntary camera rotations needed to be avoided. Finally, the personal nature of the genre (first-person educational videos) seemed to suggest that the presenter needed to employ some common social communication gestures while talking to help lead the viewer along.

⁴ <https://theta360.com/en/about/theta/v.html>

It was decided to attempt to bring together some of these thoughts into an approach that might best convey the messages the presenters were trying to get across, while maintaining some intimacy with the viewer. By exploring some variations on different shots, the hope was to develop some new shots to assist in guiding people making such videos, and to provide some examples of these.

Preliminary Work in Christchurch

The videos I captured in Christchurch are summarised in Table 2, which also contains the lessons learned from viewing them in a VR viewer, and links to the individual videos. Each of these videos can be viewed using a standard Web browser using the link information provided. The mouse can be used to move the viewpoint around while a video is playing, simulating the head movement when the VR viewer is used.

Table 2: List of test videos recorded in Christchurch.

Description	Lessons Learned/Confirmed	YouTube Video Link
Colin Monteath, describing his library. He is using a selfie-stick and trying to keep the camera at eye level.	<ul style="list-style-type: none"> Keep the camera at eye level Camera resolution might not be high enough to capture small details (e.g., the books) 	https://youtu.be/L0vbJ3GYU_Y
Colin Monteath, describing his book shop. He is using a selfie-stick and trying to keep the camera at eye level.	<ul style="list-style-type: none"> Do not rotate around too much Using a “follow me” movement works better than the presenter pulling the viewer along Training the presenter can help a lot 	https://youtu.be/EBULOFg8qXg
Paul Broady, walking around the UC campus, describing lichen & algae.	<ul style="list-style-type: none"> Camera rotation should be minimised Training the presenter can help a lot 	https://youtu.be/GK2RaNh8iqU
Paul Broady, on the UC campus, walking into a stream to describe algae.	<ul style="list-style-type: none"> Things the viewer needs to see close-up should be moved towards the camera 	https://youtu.be/N6LOA5GzZK8
Paul Broady, on the UC campus, walking next to a stream describing mites.	<ul style="list-style-type: none"> Showing close-up content to the camera can be effective. 	https://youtu.be/2KJyeI4Ct0Y
Paul Broady, describing moss inside a hot house on the UC campus.	<ul style="list-style-type: none"> A tripod can be effective for content that is fixed or far away 	https://youtu.be/79dcXi3bTnU

Follow-on Work in Antarctica

Following the test shoots, the lessons learned from them were then applied when deciding the methods used for capturing once in Antarctica. I decided to use any opportunities I could to capture content, and then to work through them once back in Christchurch in preparation for this report. Table 3 shows the list of content that was captured in

Antarctica, along with the further lessons learned or confirmed from them. As before, links are provided that show the captured content.

Table 3: List of videos recorded in Antarctica.

Description	Lessons Learned/Confirmed	YouTube Video Link
Chris Long, describing the area around Room with a View.	<ul style="list-style-type: none"> Far objects (e.g., Mt. Erebus) can be effectively indicated using pointing (and looking by the viewer). 	https://youtu.be/Szexa2gabHQ
Emma Beech, describing the history around Discovery Hut.	<ul style="list-style-type: none"> Lighting control is important Low camera resolution will not capture details well. 	https://youtu.be/iZ9ot4E3ddU
Jonathan Taylor, describing the scene around Hut Point.	<ul style="list-style-type: none"> Scenes with distant content can be effectively captured with a third person holding the selfie stick. 	https://youtu.be/jTUeHxrTLrk
Chris Long, describing the layout of the TAE Hut, before going inside.	<ul style="list-style-type: none"> 4k resolution (as opposed to the 1080 resolution used until this one) shows much better detail. 	https://youtu.be/Pkr_WCkQU9E
Chris Long, describing the inside of the TAE Hut, followed by me walking around the various rooms, pausing so the viewer can look around.	<ul style="list-style-type: none"> Pausing to let the viewer explore with their head can be effective. 4k resolution makes a difference 	https://youtu.be/FpIsaCrkfNo
Several members riding along in a Hagglund.	<ul style="list-style-type: none"> A centred, eye-level tripod can be an effective viewpoint for capturing action of multiple actors 	https://youtu.be/COH61duyCkE

Variables to Consider when Planning 360-degree Video Content

From these two sets of captured videos, we can identify several aspects to consider when planning content for capture and consumption using 360-degree video. They involve the nature of the presenter, the content to be presented, and the way the content will be viewed.

Content Capture

There are several variables to consider before capturing a SMEs discussing their work. These include:

- 1) **Posture:** Is the SME standing, seated in a chair, or walking around?
- 2) **Eye Level:** Is the SME looking up at, down on, or even with the camera (and hence the future viewer)?
- 3) **Viewpoint Mobility:** Is the camera attached to a fixed tripod, or is it under the control of the SME or third party (e.g., on selfie stick)?

- 4) **Subject-matter Proximity:** Will the SME be referring to content that is close-by (e.g., lichen on a rock) or far away (e.g., a distant volcano, penguins)?

Content Consumption

In addition, there are several variables to consider when choosing possible options for viewers of the content. These include:

- 1) **Posture:** Is the viewer standing, seated in a chair, or seated on the floor?
- 2) **Available Technology:** What technology is available to the viewer (e.g., phone/tablet, desktop, VR headset)?
- 3) **Group/Solo Viewing:** Is the viewer alone, or is the content to be viewed by a group?

Intuitively, many of these variables will have some interdependency. For example, if the SME is standing during capture, it may be better for the viewer to also be standing, and if group viewing is needed, a VR headset may not be the best option.

Possible Capture Scenarios

These aspects can be combined in numerous ways, and Table 4 lists the combinations that were explored in this work.






Table 4: Capture/consumption variables explored in this work

SME Posture	SME Eye Level	SME Mobility	SME Proximity	Scenario
Standing	Level	Mobile	Far	Chris Long describing the view from Room with a View in Antarctica.
Standing	Level	Mobile	Near	Chris Long describing TAE Hut on the Ice.
Standing	Level	Mobile	Near	Emma Beech describing Discovery Hut in Antarctica.
Chair	Level	Static	Near	Chris Long driving the Hagglund from Windless Bight in Antarctica.
Standing	Level	Mobile	Close	Colin Monteath, showing some books/artefacts at his home
Standing	Level	Static	Close	Paul Broady, showing moss in a bed at UC
Standing	Level	Mobile	Close	Paul Broady, showing moss and river algae at UC.

Recommendations

The exploration of 360-degree video capture options in this exercise has led me to define several new shot types (Table 5), and a set of guidelines. As with the traditional shot types listed in Table 1, these are designed to support the creative process of effective storytelling given the strengths and limitations of the given medium.

Table 5: Proposed 360-degree Video Camera Shot Types
(each row of images should be viewed as a sequence, from left to right)

#	Shot Name	Description
1	Conversation Shot	A shot that keeps the camera at a normal distance for having a one-on-one conversation. 
2	Come Along Shot	A camera movement that seems to beckon the viewer to move along with the presenter at a measured pace. 
3	Look There Shot	A hand and head gesture where the presenter points and looks out at a distant point of interest, then looks back at the camera. 
4	Look Here Shot	A camera movement that brings the camera towards a very close, immovable point of interest. 
5	Show Me Shot	A hand movement where an object is brought close to the camera by the presenter. 
6	Pause Shot	Keeping the camera stationary for some time to let the viewer look around on their own.
7	Inside-out Shot	The camera is placed inside a circle of people.

The following list is a set of guidelines to help 360-degree filmmakers better prepare for capturing content:

- Presenters should
 - be trained on how to present using this medium,
 - speak to the camera as if it were a live listener, and
 - use common communication gestures (pointing, showing things) to the camera while presenting.
- If the presenter needs to show something like a hand-held object to the viewer, then this should be done in the same way (i.e., at the same distance) as if the camera were a real person.
- If the camera is held by the presenter (e.g., on a selfie stick), then
 - it can be moved along with the presenter, as long as it is moved at a similar pace to walking along with a person, and
 - it should not be rotated around the presenter, as this creates large rotation movement of the view, and will cause motion sickness in the viewer.
- The camera should be kept at eye level with the presenter.

It is hoped that these shots and guidelines can be used as a starting point for developing a lexicon of how to describe effective content creation using 360-degree video.

References

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science*, 323(5910), pp. 66–69. doi: 10.1126/science.1167311

Slater, M., Usoh, M. and Steed, A. (1994). Depth of Presence in Virtual Environments, *Presence: Teleoperators and Virtual Environments*, 3.2, pp. 130- 144.